

*Ensayos–Volumen XXVII, núm.1, mayo 2008, pp. 1-32*

## **¿Es la política social una causa de la informalidad en México?<sup>1</sup>**

Gerardo Esquivel Hernández\*  
Juan Luis Ordaz-Díaz\*\*

### **Resumen**

Levy (2008) ha sugerido que la expansión de programas sociales no focalizados, podría inducir un aumento en la informalidad económica en países como México. Sin embargo, esta hipótesis parte del supuesto de la existencia de mercados laborales integrados y competitivos en una economía. En este trabajo, se hace una prueba empírica de dicho supuesto para el caso de México y, se encuentra evidencia de que existe un premio salarial en el mercado laboral formal, es decir, que individuos con características similares ganan más cuando trabajan en el sector formal de la economía que cuando lo hacen en la informalidad. En ese sentido, se concluye que dado que el mercado laboral en México está segmentado, por lo tanto, el aumento en los programas sociales no estaría dando lugar a aumentos en la tasa de informalidad en el país.

**Palabras clave:** *sector Informal, economía dual, política social.*

### **Abstract**

Levy (2008) has suggested that the expansion of non-targeted social programs could be inducing an increase in economic informality in countries like Mexico. This hypothesis, however, assumes the existence of integrated and competitive labor markets in the economy. In this work we test this assumption for the Mexican case and we find that there is a wage premium in Mexico's formal labor market. This means that an individual earns a higher wage when she works in the formal sector of the economy than an individual with similar characteristics that works at the informal sector. We therefore conclude that Mexico's labor market is segmented and that an increase in social programs is not causing an increase in informality in the country.

**Key words:** *informal sector, dual economy, social policy.*

**Clasificación JEL:** *J21, J31, O17*

---

\* El Colegio de México. Correo electrónico: [gesquive@colmex.mx](mailto:gesquive@colmex.mx)

\*\* CEPAL

<sup>1</sup> Se agradecen los comentarios de dos dictaminadores anónimos y de Juan Carlos Moreno Brid. Cualquier error es responsabilidad de los autores.

## Introducción

Muchas economías en desarrollo se caracterizan por contar con un amplio sector económico que está constituido por empresas que no pagan impuestos, y tampoco están sujetas a la regulación gubernamental; así como por trabajadores que no reciben prestaciones laborales, ni pagan impuestos por los ingresos recibidos. Este segmento de la economía es conocido como “el sector informal”, cuya amplia presencia es característica de muchos países de América Latina, inclusive de México.<sup>2</sup>

Existen dos grandes vertientes teóricas que han tratado de explicar la presencia del sector informal en una economía. Por un lado, se argumenta que el sector informal es el resultado de decisiones óptimas por parte de empleados y empleadores, los cuales optan por desempeñarse voluntariamente en el sector informal de la economía, debido a que éste les ofrece algunas ventajas, como: una mayor flexibilidad laboral, o la capacidad de evadir el pago de impuestos (Amaral y Quintín [2006], Galiani y Weinschelbaum [2008] y Levy [2008]). Por otro lado, existen algunos estudios basados en la tradición del modelo clásico de Harris y Todaro (1970), los cuales sugieren que la presencia del sector informal se debe a la existencia de algunas imperfecciones de mercado, que podrían dar lugar a un racionamiento en el mercado laboral, así como a la existencia de una economía dual (Rauch, 1991). Esta visión implica que la existencia de un sector informal en una economía, no es necesariamente el resultado de elecciones óptimas por parte de empresas y trabajadores.

De acuerdo con el primer enfoque, Levy (2008) y otros autores han sugerido que en equilibrio, los trabajadores del sector informal deberían recibir un *premio salarial*; es decir, que los trabajadores informales percibirían un salario más elevado que el de los individuos con características similares que trabajan en el sector formal, ya que eso les permitiría compensar la ausencia de las prestaciones sociales o laborales que si estarían recibiendo los trabajadores del sector formal, como el acceso a la seguridad social, por ejemplo. Más aún, Levy (2008) ha sugerido que la presencia de un creciente número de programas sociales con amplia cobertura, podría promover mayor informalidad en una economía, al gravar el empleo asalariado formal y subsidiar al empleo informal mediante los programas de protección social. Así, los trabajadores tendrían incentivos cada vez mayores para desempeñarse en el sector informal, ya que de cualquier forma estarían recibiendo algunas prestaciones sociales como

---

<sup>2</sup> Existen definiciones alternativas de lo que es el sector informal de la economía. Para una discusión sobre este concepto y sobre su medición en el caso de América Latina, véase Perry et al. (2007), así como Gasparini y Tornarolli (2007).

### *¿Es la política social una causa de la informalidad en México? 3*

acceso a guarderías, pensiones mínimas, servicios de salud, entre otras. En consecuencia, la presencia cada vez más difundida de algunos programas sociales podría dañar la productividad y el crecimiento económico, sobre todo si los programas sociales se expanden rápidamente como ha ocurrido recientemente en el caso de México (Levy, 2008).

Este punto de vista sobre la economía se contrapone con la perspectiva implícita en los modelos que enfatizan la existencia de una economía y un mercado laboral de carácter dual, en donde el sector informal es un sector residual de la economía. Según este enfoque, precisamente por tratarse de un sector residual o secundario, los salarios y las condiciones laborales del sector informal serían inferiores a las que se podría obtener en el mercado formal y, por lo tanto, individuos con características similares recibirían un salario menor en el sector informal que en el formal; es decir, existiría un *premio salarial* para los trabajadores asalariados formales.

Tales perspectivas contrapuestas sobre el sector informal han sido sujeto de múltiples análisis empíricos, sin que hasta ahora se haya obtenido resultados concluyentes. Por lo tanto, la presente investigación tiene como propósito aportar mayor luz en la discusión de estos temas; para ello, se pretende responder a una interrogante fundamental: ¿están integrados o segmentados los mercados laborales de la economía formal e informal en México? Con tal finalidad, se investigará si existe un premio salarial a la informalidad, como lo sugiere el enfoque que supone que los mercados laborales son competitivos e integrados. O, por el contrario, si existe un premio salarial a la formalidad, como lo sugieren los modelos que enfatizan el aspecto dual de una economía. Si ocurriese lo primero, entonces es posible suponer que un aumento en los programas sociales podría dar lugar a un aumento en la informalidad; mientras que si ocurriese lo segundo, la falta de integración de los mercados laborales sugeriría que la política social no es la causa de una mayor informalidad en la economía.

Para estimar el efecto del tipo de sector (formal o informal), en la brecha salarial entre sectores, se hará uso de técnicas semiparamétricas basadas en la metodología del “*propensity score matching*” (o emparejamiento por propensión).

La utilización de esta técnica permite comparar los ingresos de individuos similares en ciertas características, que trabajan en uno u otro sector de la economía. De esta forma, es posible identificar si, por el simple hecho de trabajar en uno u otro sector, un individuo gana más o menos que un individuo similar que labora en otro sector. Esta metodología, muy utilizada para propósitos de evaluación, nos permitirá estimar lo que se

#### 4 Ensayos

conoce como un “efecto de tratamiento”, en donde el tratamiento en este caso será el de laborar en el sector formal de la economía.<sup>3</sup> Para que sea factible una evaluación con mayor precisión sobre las tendencias de la informalidad en el país y sobre la evolución de la brecha salarial entre sectores, el estudio se realizará para un periodo relativamente largo (1995-2004), el cual coincide con la expansión acelerada que se ha observado en algunos programas sociales no focalizados.

La estructura de este trabajo después de la introducción, es la siguiente. La sección uno (I) describe los niveles de informalidad en México y ofrece una síntesis de los resultados obtenidos en estudios previos. La sección dos (II) contiene la hipótesis que señala que los programas sociales fomentan la informalidad. La sección tres (III) expone la metodología y los datos empleados. La sección cuatro (IV) muestra los resultados de la estimación de la brecha salarial entre sectores. Se finaliza con un resumen y las principales conclusiones.

#### **I. La informalidad en México: ¿exclusión o escape?**

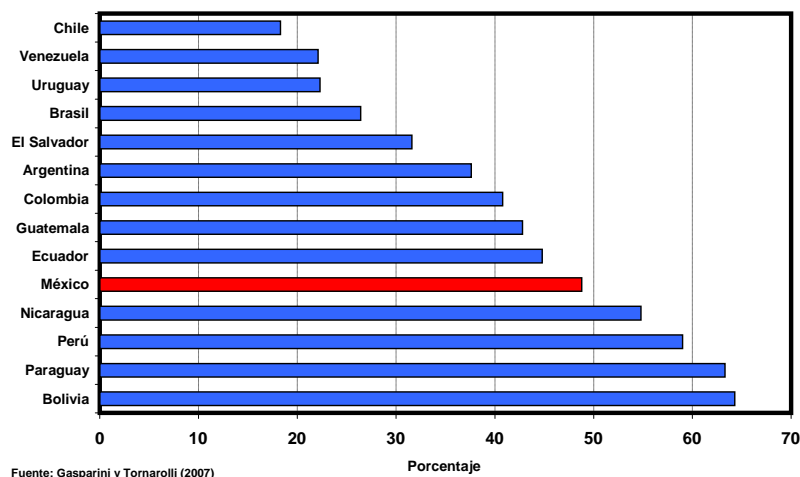
Como ya se mencionó antes, muchos países en desarrollo se caracterizan porque un amplio sector de su actividad económica tiene lugar dentro del ámbito denominado “informal”. Esta situación también es característica de muchos países de América Latina, y México no es la excepción (Gasparini y Tornarolli, 2007). Esta situación se puede observar en la gráfica 1, en donde se muestra el porcentaje de los trabajadores urbanos que son informales en varios países de América Latina.

Estos datos utilizan un criterio de ausencia de prestaciones sociales y laborales para definir la informalidad y se basan en encuestas de hogares, que fueron realizadas alrededor del año 2003. En dicha gráfica, se observa que México presenta una tasa de informalidad por encima del promedio de América Latina y muy por encima de la que se observa en países de ingresos similares como Chile, Uruguay, Brasil o Argentina.

---

<sup>3</sup> Esta metodología ha sido aplicada en un contexto similar por Pratap y Quintin (2006) y Levy (2008), cap.5.

Gráfica 1. Tasa de Informalidad en Áreas Urbanas en algunos países de América Latina, circa, 2003



Analizar y comprender los factores que determinan si un trabajador lo hace en el sector formal o informal es relevante ya que a partir de ello se pueden derivar diferentes implicaciones de política económica. Así, de acuerdo con Perry et al. (2007), una gran parte del debate sobre la naturaleza del sector informal se puede dividir en dos enfoques. El primero, denominado el enfoque de la *exclusión*, sugiere que los trabajadores informales forman parte del sector en desventaja de un mercado de trabajo segmentado, que surge del carácter dual de una economía y de ciertas rigideces institucionales. Esto implica que la participación en el sector informal no sería necesariamente el resultado de una elección óptima por parte de empresas y trabajadores, sino que sería originado por algunas imperfecciones del mercado que podrían dar lugar a un cierto racionamiento en el mercado laboral. Lo anterior supone la existencia de mercados laborales no competitivos, sin libre transición entre sectores -por ejemplo, entre los sectores formal e informal- y con racionamiento de empleo.<sup>4</sup>

Alternativamente, existe el enfoque denominado de *escape*, el cual señala que los trabajadores pueden encontrar condiciones similares en ambos sectores pero que la participación en el sector informal se explica por una economía con un Estado de Derecho relativamente débil y una carga regulatoria o fiscal relativamente importante.<sup>5</sup> De acuerdo con este enfoque,

<sup>4</sup> Este es el enfoque clásico de Harris y Todaro (1970) y que fuera modelado por Rausch (1991).

<sup>5</sup> Esta perspectiva está inspirada en el trabajo de Hernando de Soto (1987).

## 6 Ensayos

los trabajadores deciden de forma voluntaria incorporarse al sector informal debido a que este sector les ofrece algunas ventajas como pueden ser una mayor flexibilidad laboral o la capacidad de evadir el pago de impuestos. Una implicación de esta visión del mercado de trabajo es que lo observado en el mercado laboral es el resultado de decisiones óptimas y, por ende, implicaría la existencia de un equilibrio en dicho mercado.

Estas visiones contrapuestas sobre las características del mercado laboral en presencia de un sector formal y uno informal, han sido analizadas en diferentes estudios empíricos. En este contexto, el caso de México ha sido uno de los más estudiados y, paradójicamente, es uno de los que ha generado los resultados más contradictorios sobre estas dos visiones del mercado laboral. Por ejemplo, Juárez (2006) concluye que el mercado laboral mexicano está segmentado. Así, al analizar las diferencias salariales entre trabajadores asalariados formales e informales en México, para el primer y segundo trimestre de 1996, y con base en regresiones salariales, la autora encuentra que los trabajadores del sector formal ganan, en promedio, más que los informales, incluso una vez que se ha controlado por características observables.

Por otro lado, Gong y van Soest (2002) encuentran que las dos visiones sobre el mercado son válidas, y que todo depende del nivel de educación de los trabajadores. De esta forma, con datos de 1992 y 1993, estos autores estiman un modelo de panel con efectos dinámicos aleatorios y encontraron que los rendimientos a la educación son más bajos en el sector informal y que en ambos sectores los salarios aumentan con el nivel educativo. Además, estos autores encuentran diferencias salariales positivas entre el sector formal e informal para los más educados y diferencias salariales pequeñas (o incluso negativas) para los de menores niveles educativos. Así, para estos autores, los mercados laborales para trabajadores poco calificados se encuentran integrados, pero no ocurre lo mismo en el caso de trabajadores con mejor nivel de calificación.

En cambio, algunos estudios han validado la visión del sector informal como una opción de *escape*, al sugerir que el mercado laboral mexicano es de alta movilidad y, competitivo. Incluso, algunos estudios han encontrado que existe un premio a la informalidad, lo que implica que los trabajadores informales ganarían más que individuos con características similares y que trabajan en el sector formal de la economía. Por ejemplo, Maloney (1999) analiza los patrones de movilidad de los trabajadores mexicanos en distintos sectores laborales durante el periodo 1990-92, y caracteriza las interacciones entre los distintos sectores laborales, cuando los compara con los patrones implícitos que predice la visión estándar de mercados segmentados. Maloney concluye que los individuos que pasan del sector formal al informal incrementan sus ingresos, en promedio, en un veinticinco por ciento (25%);

### *¿Es la política social una causa de la informalidad en México? 7*

mientras que los que realizan la transición opuesta, sufren una disminución en sus ingresos.

En esta misma línea, Navarro y Schrimpf (2004) desarrollan un modelo dinámico de dos sectores, en el que cada persona elige el sector en el que opta trabajar, a través de un proceso de maximización de la utilidad. Empleando métodos para modelos de elección discreta, concluyen que no existe segmentación en el mercado de trabajo mexicano en el periodo 1997-1998, ya que sus resultados muestran que las personas escogen el sector en el que trabajan, con base en las mismas características que determinan su nivel de escolaridad. Por otro lado, Marcouiller, Ruiz de Castilla y Woodruff (1997), utilizando métodos de regresión corregidos por autoselección buscan responder si los trabajadores del sector informal ganan menos que trabajadores de características semejantes del sector formal en México, El Salvador y Perú. Para el caso de México, estos autores encuentran que existe un premio de ingresos asociado al sector informal de la economía.

Finalmente, Perry et al. (2007), quienes resumen la evidencia empírica para muchos países de América Latina y para el caso de México, concluyen que la evidencia de segmentación laboral es débil y que la mayoría del sector informal puede considerarse como “voluntario”. En esta misma línea de argumentación, Levy (2007) supone que ante la falta de pagos y prestaciones laborales que reciben los trabajadores del sector formal, los trabajadores informales podrían estar recibiendo un premio que les permita compensar la ausencia de tales beneficios. Esta situación, según él, puede tener implicaciones sobre cómo influyen los programas sociales en la dinámica del mercado laboral. La siguiente sección está orientada a ilustrar la lógica de este argumento.

## **II. Hipótesis de la relación entre política social e informalidad<sup>6</sup>**

Considérese una economía compuesta por individuos que se diferencian en términos de sus características individuales ( $X$ ), que se emplean ya sea en el sector formal ( $F$ ) o en el informal ( $I$ ), y dependiendo de en qué sector se emplean, los individuos pueden percibir un salario formal dado por  $w_f(X)$ , o un salario informal definido por  $w_i(X)$ , los cuales están en función de las características personales.

En mercados laborales competitivos, con libre transición entre sectores y sin seguridad ni protección social, el valor del producto marginal y los

---

<sup>6</sup> Esta sección está basada en Levy (2007, 2008).

## 8 Ensayos

salarios para individuos con características similares serían equivalentes en ambos sectores, es decir,  $w_f = w_i$ .

Sin embargo, los trabajadores del sector formal reciben otras prestaciones tales como el derecho a una pensión o los beneficios de la seguridad social. Estos beneficios tienen como característica que son pagados con la participación de los trabajadores y las empresas del sector formal. Al mismo tiempo, los trabajadores del sector informal pueden beneficiarse de los programas de protección social del gobierno, los cuales no tienen una relación directa con el trabajo. Este tipo de programas son cubiertos con recursos de los impuestos generales y no necesariamente con las contribuciones de los trabajadores o las empresas del sector informal. Por lo tanto, en un mercado competitivo, los ingresos de los trabajadores informales dependen no sólo de sus propios salarios sino, también, de los beneficios que reciben los trabajadores del sector formal.

Para ilustrar lo anterior, considérese un mercado de trabajo competitivo en el que los costos de los beneficios de la seguridad social están dados por  $T_f$ . En este caso, los costos para una empresa de contratar a un trabajador en el sector formal serán  $(w_f + T_f)$ . Además, supóngase que  $\beta_f \in [0,1]$  es el coeficiente de la valoración personal sobre los beneficios de la seguridad social. Así, desde el punto de vista de un trabajador, la utilidad de un empleo formal será  $w_f + \beta_f T_f$ . Si los trabajadores valoran la seguridad social plenamente, es decir, si  $\beta_f = 1$ , los costos de contratación de un trabajador en el sector formal serán iguales a la utilidad del mismo. En cambio, si los trabajadores no valoran plenamente los beneficios de la seguridad social, el costo de contratar a un empleado formal sería superior a la utilidad que éste recibe, es decir,  $(w_f + \beta_f T_f) < (w_f + T_f)$ . Nótese que la diferencia entre estas dos expresiones es equivalente a un impuesto sobre el empleo asalariado formal.

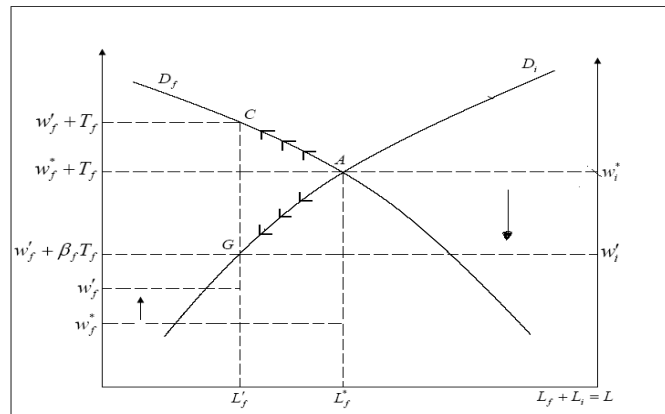
Por otra parte, en el sector informal no se tienen costos de la seguridad social, pero los trabajadores obtienen beneficios de los programas sociales, cuyo costo por trabajador es  $T_i$  y los cuales son valorados por los trabajadores en  $\beta_i T_i$ , con  $\beta_i \in [0,1]$ . En este caso,  $T_i$  no es pagado por los trabajadores o empresas. Dado lo anterior, el equilibrio en la economía estará representado por:



$$\begin{aligned} p \cdot PMg_f &= (w_f + T_f) \\ p \cdot PMg_i &= w_i \\ w_i + \beta_i T_i &= w_f + \beta_f T_f \\ L_i + L_f &= L \end{aligned} \tag{2.1}$$

Las primeras dos ecuaciones ilustran la igualdad entre el valor del producto marginal ( $p \cdot PMg$ ) y el costo por trabajador en cada uno de los sectores, la tercera ecuación ilustra el equilibrio entre sectores desde la perspectiva de los trabajadores (es decir, la igualdad de utilidades en ambos sectores), mientras que la última ecuación indica que, en equilibrio, la suma de los trabajadores en ambos sectores ( $L_i$  y  $L_f$ ) debe ser igual a la fuerza laboral total ( $L$ ).

A partir de (2.1), supóngase que inicialmente la seguridad social es valorada plenamente y que no hay programas de protección social (es decir,  $\beta_f = 1$  y  $T_i = 0$ ). La gráfica 2 representa esta situación, en donde  $D_f$  y  $D_i$  representan las demandas de trabajo formal e informal, respectivamente. En este caso, el equilibrio de los mercados laborales se encuentra en el punto A, donde se tienen  $L_f$  trabajadores en el sector formal y el resto ( $L - L_f$ ) en el sector informal. En esta solución, el salario en el sector formal será  $w_f^*$ , la utilidad para los trabajadores del sector formal es  $w_f^* + T_f$ , mientras que el salario en el sector informal está dado por  $w_i^* = w_f^* + T_f$ . Nótese que estos resultados implican un premio salarial a la informalidad, es decir,  $w_i^* > w_f^*$ .

**Gráfica 2. Seguridad social con sector formal e informal**

Fuente: Levy (2007)

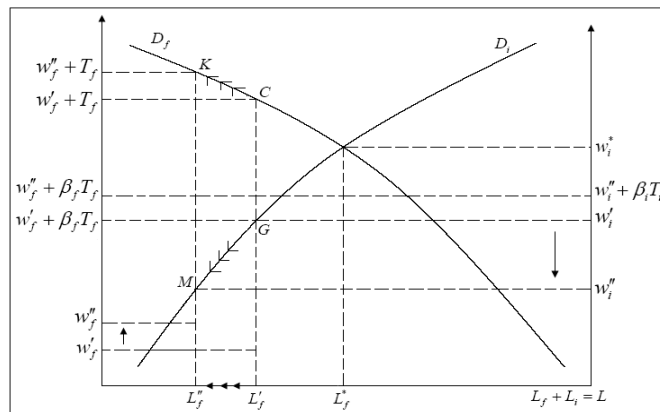
Ahora, suponga que la seguridad social no es valorada plenamente, o lo que es lo mismo  $\beta_f < 1$ . En este caso, entonces  $w_i^* > w_f' + \beta_f T_f$ , por lo tanto, los trabajadores del sector informal estarían obteniendo una mayor utilidad que los del sector formal. Como no existen restricciones a la movilidad, esto hará que algunos trabajadores formales se trasladen hacia el sector informal de la economía. Al haber menos trabajadores en el sector formal, los salarios en este sector aumentarán -por ejemplo, a  $w_f'$ - y las empresas formales se moverán de A a C (gráfica 2). Con más trabajadores en el sector informal, los salarios en este sector caerán a  $w_i'$ ; mientras que en el sector informal, se incrementará el empleo de A a G. En cualquier caso, en el nuevo equilibrio, a menos que  $\beta_f = 0$ , seguirá existiendo un premio salarial a la informalidad, ya que  $w_i' > w_f'$ .

Ahora, introduzcamos en el análisis la existencia de programas de protección social, de los cuales pueden beneficiarse los trabajadores del sector informal ( $T_i > 0$ ). En este caso, una vez que se resuelve (2.1), se tiene que en los puntos G y C de la gráfica 3, los trabajadores en el sector formal obtendrían una utilidad dada por  $w_f' + \beta_f T_f$ ; mientras que los del sector informal, estarían recibiendo  $w_i' + \beta_i T_i > w_f' + \beta_f T_f$ ; con lo cual, los trabajadores del sector informal estarían mejor que los trabajadores del sector formal. Así, estos trabajadores tendrán un incentivo para moverse

### ¿Es la política social una causa de la informalidad en México? 11

hacia el sector informal y, ante una menor oferta laboral en el sector formal, los salarios en este sector se incrementarán de  $w_f'$  a  $w_f''$ ; así, las empresas se habrán movido al punto  $K$ . Por su parte, los salarios del sector informal caerán de  $w_i'$  a  $w_i''$ . En ese sentido, la presencia de programas de protección social habría inducido un aumento en la informalidad. Nótese que en tal caso, al igual que en los anteriores, la libre movilidad entre sectores supone la existencia de un premio salarial a la informalidad, ya que  $w_i'' > w_f''$ .<sup>7</sup>

**Gráfica 3. Mercado de trabajo con seguridad y protección social**



Fuente: Levy (2007)

En resumen, debe notarse que para que los programas sociales puedan causar una mayor informalidad en la economía, es necesario que no existan restricciones a la movilidad entre sectores y que, en términos salariales, exista un premio a la informalidad. Congruente con ello, el propósito fundamental de la presente investigación será analizar, empíricamente, si

<sup>7</sup> La relación salarial entre ambos sectores depende del nexo entre  $\beta_f T_f$  y  $\beta_i T_i$  (ecuación 2.1). Aunque, considerando que algunas estimaciones de  $T_f$  indican que su valor es más de cuatro veces el valor de  $T_i$  (Levy, 2008), entonces  $\beta_i$  debería ser superior a cuatro veces el valor de  $\beta_f$  para que no hubiera un premio salarial a la informalidad, lo cual se considera muy poco plausible. Ello, sin considerar que los trabajadores formales también podrían beneficiarse de algunos programas de protección social, lo cual haría esta situación aún menos probable.

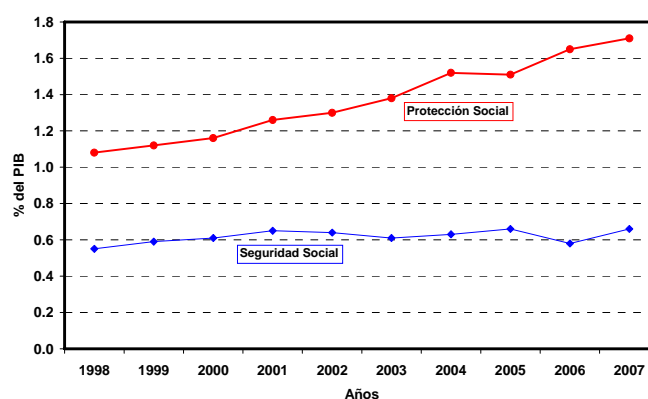
## 12 Ensayos

esta condición se cumple en el caso mexicano; lo cual permitirá analizar si los programas sociales en México tienden o no a fomentar la informalidad.

### *Gasto en protección social en México*

El motivo de la creciente preocupación con respecto al efecto que pudiera tener el gasto en protección social sobre la informalidad, deriva de que recientemente se ha dado un crecimiento muy importante en los recursos públicos destinados a este tipo de programas. Lo anterior es cierto no sólo en términos absolutos, sino incluso si se compara con los recursos destinados a la seguridad social -excluyendo al programa Oportunidades- o, con respecto al PIB. En efecto, a lo largo del periodo 1998-2007, el monto de recursos públicos destinados a programas de protección social creció en 110%, mientras que los recursos destinados a la seguridad social, lo hicieron en alrededor de 40%. Como porcentaje del PIB, los primeros pasaron de 1.1% a 1.7%, mientras que los segundos se mantuvieron relativamente constantes, en alrededor del 0.6% (véase la gráfica 4).

Gráfica 4. Recursos Públicos Destinados a Programa Sociales  
(porcentaje del PIB)



Fuente: Levy (2008), Cuadro 1-7.

En este sentido, si la política social realmente tendiera a fomentar la informalidad como lo sugiere Levy (2008), se esperaría observar lo siguiente:

1) Que la informalidad hubiese aumentado en forma continua en años recientes (como resultado del aumento en los programas de protección social) y

2) Que hubiera una prima salarial a la informalidad, es decir, que  $w_i > w_f$ .

### **III. Descripción de los datos y metodología**

#### **III.1 Datos**

En este trabajo, se utiliza información de la Encuesta Nacional de Empleo Urbano (ENEU). Los datos que se usan corresponden a: las características sociodemográficas de la población urbana; la estructura ocupacional y su distribución, por sector económico; las modalidades de empleo y los ingresos de los trabajadores; y también, a las características de las unidades económicas en donde laboran los ocupados. La muestra seleccionada incluye a trabajadores remunerados de entre 12 y 65 años de edad, que fueron entrevistados en el segundo trimestre para cada uno de los años que comprende el periodo de 1995-2004. Las observaciones donde faltaba información respecto del nivel educativo, sexo o edad, no fueron consideradas. De igual forma, se eliminó de la muestra a aquellas personas que indicaron un número de horas trabajadas a la semana superior a 80, y a las que no contaban con información sobre sus salarios o para las que no era posible obtener su salario por hora. También se excluyó de la muestra a quienes se emplean por cuenta propia; ya que -en esos casos- no resulta claro qué porcentaje de los ingresos reportados, realmente corresponden a los rendimientos del capital invertido.<sup>8</sup>

Como se mencionó, no existe una forma única de definir al sector informal y generalmente se le asocia con actividades no registradas oficialmente, con autoridades fiscales o de seguridad social. En consecuencia, en este estudio, se utilizará el “criterio de la seguridad social”<sup>9</sup> y se clasificará como formal al trabajador que recibe, por lo menos, una de las prestaciones siguientes: a) incorporación al Instituto Mexicano del Seguro Social (IMSS); b) vacaciones con goce de sueldo; c) participación de utilidades; d) afiliación al Instituto de Seguridad y Servicios Sociales de los

---

<sup>8</sup> Incluir a los auto-empleados como parte del sector informal, tal como se ha hecho en varios de los trabajos previos, puede llevar a obtener resultados imprecisos cuando se analiza la segmentación laboral, debido a que los ingresos de estos trabajadores usualmente incluyen los rendimientos del capital. Así, lo que ellos reportan como salario puede incluir un porcentaje que en realidad corresponde a los retornos del capital invertido. Por lo tanto, cuando se incluye a este grupo de trabajadores en un análisis de este tipo se tienden a sesgar los resultados en la dirección de encontrar un premio salarial a la informalidad.

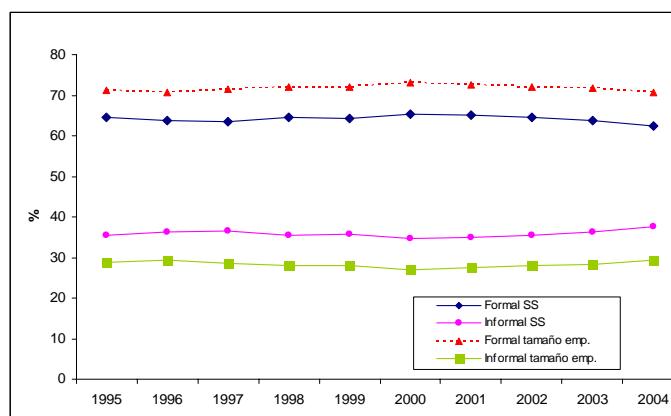
<sup>9</sup> Algunos de los trabajos previos realizados sobre este tema, utilizan el “criterio del tamaño de la empresa”, con el que un trabajador se define como informal, según el número de empleados en el establecimiento en el que labore. Esta definición, sin embargo, es cuestionable ya que no todas las pequeñas empresas son informales y no todas las empresas grandes son formales. En cualquier caso, los resultados obtenidos en este trabajo son robustos a una definición de informalidad de este tipo, y los datos están disponibles a solicitud de los interesados (véase, por ejemplo, los resultados de las gráficas 5 y 6).

#### 14 Ensayos

trabajadores del Estado (ISSSTE); e) derecho a crédito para la vivienda; f) goce de la prestación del Sistema de Ahorro para el Retiro (SAR); g) seguro médico particular o seguro de salud.<sup>10</sup> Además, y sólo para propósitos comparativos, en algunas ocasiones usaremos una definición de informal determinada por el “criterio del tamaño de la empresa”, en donde empleados en empresas con cinco o menos trabajadores son considerados como empleados informales.

La gráfica 5 muestra cómo ha evolucionado la proporción de trabajadores en ambos sectores (formal e informal) durante el periodo 1995-2004, de acuerdo con las definiciones de formalidad, antes mencionadas.

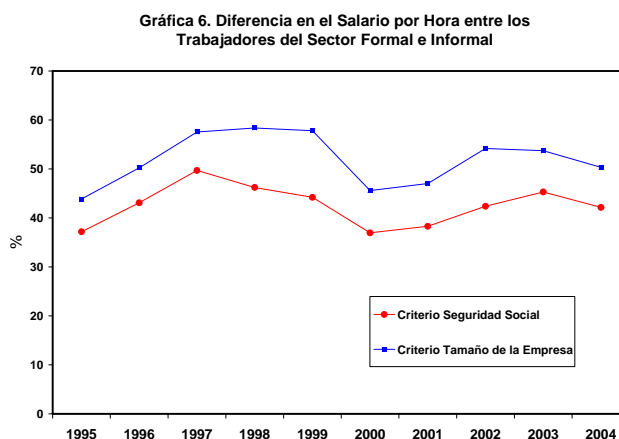
**Gráfica 5. Tamaño del sector informal en México**



Como se observa en la gráfica 5, la informalidad se ha mantenido relativamente estable a lo largo de todo el periodo, por lo que la participación en el sector informal no parece estar reaccionando ante los cambios recientes en los gastos en protección social. No obstante, este comportamiento no es una evidencia contundente de que los programas sociales no estén fomentando la informalidad en México. Para ello, se requiere probar la hipótesis de integración o segmentación de los mercados laborales, lo que se realizará a continuación. Sin embargo, antes del análisis de la segmentación o integración de los mercados laborales, mostraremos algunas características básicas de los sectores formal e informal, de acuerdo con los criterios seleccionados.

<sup>10</sup> Aunque el *Aguinaldo* es una prestación que se establece en la Ley Federal de Trabajo, no se utilizó para clasificar a los trabajadores como formales o informales, ya que existe evidencia de que muchos empleados informales también reciben esta prestación.

En primer lugar, la gráfica 6 muestra el diferencial salarial observado entre ambos sectores a lo largo del periodo, acorde con los dos métodos de clasificación mencionados. Como se muestra en esta gráfica, los trabajadores del sector formal tienden a obtener un salario que fluctúa alrededor de 40% ó 50%, por encima del salario que perciben los trabajadores del sector informal, dependiendo de la definición utilizada. Esta brecha, sin embargo, aún no controla por posibles diferencias en las características de los trabajadores de uno u otro sector, por lo que solamente debe ser considerada como informativa e ilustrativa.



Por otro lado, el cuadro 1 muestra las características individuales de los empleados de ambos sectores, para cada uno de los años del periodo a analizar.<sup>11</sup> El cuadro muestra que la proporción de mujeres es mayor en el sector formal que en el informal, que los trabajadores informales son ligeramente más jóvenes y que no hay diferencias importantes entre sectores, en el número de las horas trabajadas ni en la experiencia laboral. Sin embargo, en donde sí hay una diferencia importante entre ambos sectores es en el nivel educativo y en el tamaño de la empresa.

En lo general, la distribución de los trabajadores del sector informal tiende a concentrarse en personas con educación secundaria o inferior; mientras que la distribución de los trabajadores del sector formal, se concentra en personas con educación secundaria o superior. También, como ya se mencionó, hay un efecto asociado al tamaño de la empresa, aunque no

<sup>11</sup> Cabe señalar que la clasificación por nivel educativo considera que un trabajador tiene un nivel de Secundaria, Preparatoria o Licenciatura, si cursó por lo menos un año en el nivel correspondiente.

## 16 Ensayos

todas las empresas pequeñas son informales ni todas las empresas grandes son formales.

**Cuadro 1. Características individuales de los empleados  
(Criterio de la seguridad social)**

		1995		1996		1997		1998		1999	
		Formal	Informal	Formal	Informal	Formal	Informal	Formal	Informal	Formal	Informal
<b>Sexo</b>	Hombre (%)	62	69	62	68	62	68	62	68	62	68
	Mujer (%)	38	31	38	32	38	32	38	32	38	32
<b>Edad promedio</b>		32.8	30.9	33.0	31.4	33.0	31.3	32.8	31.6	33.2	31.9
<b>Horas trabajadas promedio</b>		43.2	42.0	43.5	42.4	43.6	42.2	43.3	42.1	43.6	43.0
<b>Experiencia laboral promedio</b>		16.8	16.9	16.9	17.2	16.8	16.9	16.6	17.3	16.9	17.6
<b>Educación</b>											
	Sin instrucción (%)	1.4	4.4	1.3	4.4	1.3	3.9	1.2	4.1	1.1	4.3
	Primaria incompleta (%)	6.2	15.1	5.8	14.3	5.8	13.2	5.6	12.9	5.3	12.9
	Primaria completa (%)	17.0	23.7	17.0	23.5	16.4	22.1	15.7	22.6	15.7	22.6
	Secundaria, Prevocacional o Equiv. (%)	38.7	33.1	37.2	33.1	37.4	34.4	37.9	35.4	37.3	34.7
	Preparatoria, Vocacional o equiv. (%)	14.8	11.9	15.5	11.8	15.7	13.1	16.1	12.6	16.7	12.6
	Licenciatura (%)	21.9	11.9	23.2	12.9	23.4	13.3	23.4	12.4	24.0	12.9
<b>Tamaño del establecimiento por número de Empleados</b>											
	5 o menos (%)	8.0	66.8	8.4	65.8	7.1	66.0	7.3	65.9	6.2	67.3
	6 a 10 (%)	8.9	14.0	8.6	14.2	8.8	14.1	8.2	13.8	8.2	13.9
	11 a 15 (%)	12.2	6.8	12.0	7.3	12.0	6.3	11.9	7.2	12.5	7.3
	16 a 50 (%)	7.9	2.1	7.2	2.5	7.1	2.5	7.3	2.4	7.7	2.2
	51 a 100 (%)	6.0	1.1	6.0	1.0	6.2	1.2	5.4	1.1	5.5	1.3
	101 y más (%)	57.1	9.2	57.9	9.2	58.8	9.9	59.9	9.6	59.9	8.0

**Cuadro 1. Continuación**

		2000		2001		2002		2003		2004	
		Formal	Informal	Formal	Informal	Formal	Informal	Formal	Informal	Formal	Informal
<b>Sexo</b>	Hombre (%)	62	66	62	67	61	67	62	67	61	65
	Mujer (%)	38	34	38	33	39	33	38	33	39	35
<b>Edad promedio</b>		33.4	32.2	33.9	32.5	34.3	33.1	34.6	33.0	35.0	33.0
<b>Horas trabajadas promedio</b>		43.5	42.3	43.5	42.4	44.3	43.9	43.7	43.2	44.3	42.3
<b>Experiencia laboral promedio</b>		17.0	17.8	17.4	17.9	17.8	18.4	17.9	18.4	18.2	18.3
<b>Educación</b>											
	Sin instrucción (%)	1.1	3.9	1.2	3.5	1.1	3.4	0.9	3.3	0.9	3.2
	Primaria incompleta (%)	5.5	12.1	4.9	11.8	4.5	11.4	4.1	10.7	3.9	10.2
	Primaria completa (%)	14.9	22.4	14.4	22.5	14.2	22.2	13.2	22.2	12.8	20.6
	Secundaria, Prevocacional o Equiv. (%)	37.0	35.6	36.6	35.0	36.7	34.9	35.9	35.5	35.8	36.0
	Preparatoria, Vocacional o equiv. (%)	17.2	13.1	18.1	14.0	18.2	14.4	19.5	15.5	20.3	16.2
	Licenciatura (%)	24.3	12.9	24.7	13.2	25.3	13.7	26.4	12.8	26.3	13.7
<b>Tamaño del establecimiento por número de Empleados</b>											
	5 o menos (%)	6.8	64.8	6.2	67.3	5.7	68.1	5.5	68.1	6.1	68.0
	6 a 10 (%)	7.9	14.2	7.4	13.6	7.2	12.8	7.9	12.8	8.2	12.8
	11 a 15 (%)	12.1	7.2	12.8	7.1	12.8	6.6	13.2	7.2	14.3	7.7
	16 a 50 (%)	7.6	2.7	6.9	2.4	7.5	2.1	7.8	2.5	8.2	2.3
	51 a 100 (%)	5.6	1.2	4.9	0.9	4.8	0.8	4.8	1.2	5.8	0.7
	101 y más (%)	60.0	9.9	61.7	8.7	62.0	8.6	60.8	8.2	57.4	8.5

## III.2 Metodología

Lo que se busca es identificar si, por el simple hecho de trabajar en uno u otro sector, un individuo puede ganar más o menos que otro individuo similar que labora en el otro sector. Para ello, se introducirá la estructura de contrafactual iniciada por Rubin (1974), con la que se analizan efectos de tratamiento sobre un determinado grupo de individuos, a partir de su comparación con un grupo de control, que no recibe tratamiento. En este caso, el grupo de tratamiento será el de los trabajadores en el sector formal



de la economía y el grupo de control será el de los trabajadores del sector informal.

La medida del efecto de tratamiento es el *average treatment effect* (ATE (X)), que se representará por  $\alpha$  y que nos dice cuál es el efecto promedio del tratamiento para un individuo con características  $X$  y que actualmente labora en el sector informal.

Dicho efecto será una medida del premio salarial que ofrece el sector formal, el cual puede ser positivo o negativo. Un valor positivo de  $\alpha$  indica que el sector formal otorga un premio a los trabajadores formales y, por lo tanto, que el mercado de trabajo está segmentado. Por otra parte, un valor negativo de  $\alpha$  indica un premio a la informalidad, lo cual sería consistente con la hipótesis de mercados laborales integrados. Para medir este efecto, se tomará la diferencia entre el resultado con tratamiento ( $w_f$ ) y sin tratamiento ( $w_i$ ).

El *efecto de tratamiento promedio* para los que fueron tratados, y condicional en las características  $X$ , se define de la siguiente forma:

$$\alpha = E(w_f - w_i | X, Formal = 1) = E(w_f | X, Formal = 1) - E(w_i | X, Formal = 1) \quad (3.1)$$

donde  $E$  representa el operador esperanza y  $X$  representa las características de los trabajadores. Como un individuo no puede ser observado en los dos estados en forma simultánea; es decir, en el sector formal y en el informal, al mismo tiempo; entonces, para calcular el último término de (3.1) se aplicará el supuesto que en la literatura de evaluación se conoce como *ignorabilidad del tratamiento* (*ignorability of treatment*).

Este supuesto indica, que:  $w_f$  y  $w_i$  son independientes, condicionados en las otras variables y que existe independencia de media condicional, esto es (Rosenbaum y Rubin, 1983, 1984):

$$\begin{aligned} E(w_i | X, Formal) &= E(w_i | X) \\ E(w_f | X, Formal) &= E(w_f | X) \end{aligned} \quad (3.2)$$

Tal supuesto lleva a definir al ATE (X) como:

$$\alpha = E(w_f | X, Formal = 1) - E(w_i | X, Formal = 0) \quad (3.3)$$

El parámetro  $\alpha$  entonces puede estimarse mediante métodos de regresión o con técnicas que utilizan el *propensity score*. Algunos trabajos previos han utilizado el método de regresión para estimar este efecto, aunque usualmente no han tomado en cuenta que los individuos podrían haberse autoseleccionado para trabajar en el sector formal o en el informal.<sup>12</sup> Si esto sucede, las estimaciones con este método estarían sesgadas y, por tanto, los resultados obtenidos serían incorrectos.

En general, se acepta que la asignación de los trabajadores entre los dos sectores laborales no es aleatoria. Por ello, si se utilizara un estimador tradicional no experimental para calcular el efecto de tratamiento, los resultados podrían estar sesgados. Para evitar este problema, la estimación de  $\alpha$  se hará empleando el método del *propensity score matching*, con el que se reduce el sesgo de estimación, al comparar los salarios de trabajadores formales e informales que tengan una propensión similar de ser tratados. Así, este enfoque permite controlar por la posibilidad de autoselección con base en características observables.

### *Propensity score matching*

Rosembaum y Rubin (1984) definen el *propensity score* como la probabilidad de recibir un tratamiento, dadas ciertas características preexistentes. Así, la similitud de los trabajadores se reflejará en el *propensity score*, lo que indica la probabilidad de participación de una persona determinada en el sector formal, dadas ciertas características personales. Así, el *propensity score* ( $p_i$ ) se especifica como:

$$p_i : P(\text{Formal} = 1 | X_i) = G(Z' \beta) \quad (3.4)$$

donde  $P(\cdot)$  denota la propensión de participación en el sector formal del trabajador  $i$ , dado el vector de las características individuales ( $X_i$ ).  $G(\cdot)$  es una función índice que genera la probabilidad de participación, la cual puede estar dada por la función de distribución acumulada normal o logística,  $Z$  es un vector que incluye a las variables,  $X$  y  $\beta$  son los coeficientes estimados.

Con esta definición y los supuestos previos, es posible estimar el efecto del tratamiento promedio condicional en el *propensity score*:

---

<sup>12</sup> No deben confundirse los conceptos de autoselección y de elección óptima. El primer término se refiere al hecho de que puede haber una serie de características específicas observables en los individuos que participan en un cierto mercado, sin que ello sea necesariamente el resultado de elecciones óptimas.

$$\alpha(p_i) = E(w_f | \text{Formal} = 1, p_i) - E(w_i | \text{Formal} = 0, p_i) = E(w_f - w_i | p_i) \quad (3.5)$$

Para obtener el ATE no condicional, se debe promediar a lo largo de toda la distribución de los *propensity scores*. Además, para que el ATE pueda ser estimado, Rosenbaum y Rubin (1983) señalan que se debe cumplir la proposición del balanceo:

$$\text{Formal} \perp X \mid P(.)$$

Si esta condición se cumple, las observaciones con el mismo *propensity score* deben tener la misma distribución de las características, independientemente de la condición de tratamiento. Esto implica que para un *propensity score* dado, la asignación al tratamiento es aleatoria, por lo que individuos con características similares tienen la misma probabilidad de recibir o no el tratamiento.

Para poder estimar la ecuación (3.5) y el ATE no condicional, es necesario hacer un *matching* entre observaciones con propensiones similares. Para realizar esto, Dehejia y Wahba (1999 y 2002) han propuesto cuatro diferentes métodos: *Nearest Neighbor Matching*, *Radius Matching*, *Kernel Matching* y *Stratification Matching* (Becker e Ichino, 2002). Estos cuatro métodos específicos se describen con un poco más de detalle en el anexo 1.

#### IV. Resultados

Se muestra los resultados de aplicar al caso de México durante el periodo 1995-2004, la metodología descrita arriba. El procedimiento tuvo dos etapas: en la primera, se estima una ecuación que determina la propensión de los trabajadores a participar en el sector formal; en la segunda, se calcula el *Efecto de Tratamiento Promedio* (ATE) mediante los cuatro métodos de estimación antes mencionados, y que son descritos en el anexo 1.

##### *Propensity Score*

Siguiendo la metodología expuesta en la sección anterior, en primer lugar se calcularon los *propensity scores* mediante un modelo tipo probit, es decir, se supuso que  $G(\cdot)$  en (3.4) corresponde a una distribución normal. Los cuadros 2 y 3 muestran los resultados que se obtuvieron, de estimar una ecuación en donde la variable dependiente es una variable dicotómica que toma el valor de 1, cuando el trabajador labora en el sector formal de la economía de acuerdo al criterio de la seguridad social.<sup>13</sup>

<sup>13</sup> No se utilizó el criterio del tamaño de la empresa debido a que, al comparar los salarios, este criterio podría estar reflejando el efecto del tamaño de la empresa en los salarios y no

Las variables independientes en la primera estimación fueron, las siguientes: una variable dummy *familiar*, que indica si otro miembro del hogar labora en el sector formal, el estado civil, el sexo, la edad y la escolaridad. Además, se incluyen dos variables interactivas: el producto de la escolaridad por la edad y el producto de la edad por el sexo. La inclusión de estas variables fue importante para que las estimaciones cumplieran con la condición de balanceo. Debido a que en los últimos años del análisis, esta primera especificación no satisfacía con un alto nivel de confianza la condición de balanceo, se utilizó otra especificación en la que se excluye al estado civil como variable explicativa. Esta nueva especificación sí satisfizo la condición de balanceo en los últimos años, por lo que hemos optado por reportar ambos resultados. Los resultados con esta segunda especificación se muestran en el cuadro 3.<sup>14</sup>

Cuadro 2. Estimación Probit (Primera especificación)  
Variable dependiente: FORMAL

	1995		1996		1997		1998		1999	
Variable	Coef.	t	Coef.	t	Coef.	T	Coef.	t	Coef.	t
Familiar										
(Formal=1)	<b>0.269</b>	20.3	<b>0.271</b>	21.4	<b>0.266</b>	21.5	<b>0.325</b>	25.0	<b>0.298</b>	22.9
Estado civil										
(Casado=1)	<b>0.093</b>	8.2	<b>0.129</b>	11.8	<b>0.107</b>	10.1	<b>0.107</b>	10.5	<b>0.102</b>	10.6
Sexo										
(Hombre=1)	<b>-0.381</b>	-13.3	<b>-0.351</b>	-12.7	<b>-0.218</b>	-8.2	<b>-0.259</b>	-9.9	<b>-0.207</b>	-8.4
Edad	<b>0.108</b>	28.2	<b>0.105</b>	28.2	<b>0.113</b>	31.1	<b>0.106</b>	30.2	<b>0.100</b>	29.7
Edad <sup>2</sup>	<b>-0.001</b>	-29.9	<b>-0.001</b>	-30.8	<b>-0.001</b>	-33.4	<b>-0.001</b>	-33.3	<b>-0.001</b>	-31.4
Escolaridad	<b>0.230</b>	20.4	<b>0.230</b>	20.8	<b>0.232</b>	21.4	<b>0.219</b>	20.6	<b>0.236</b>	23.2
Escolar <sup>2</sup>	<b>-0.006</b>	-17.7	<b>-0.006</b>	-18.8	<b>-0.006</b>	-19.4	<b>-0.005</b>	-17.5	<b>-0.006</b>	-19.1
Escoedad	<b>-0.002</b>	-6.3	<b>-0.002</b>	-6.5	<b>-0.002</b>	-6.7	<b>-0.002</b>	-6.0	<b>-0.002</b>	-7.4
Escoedad <sup>2</sup>	<b>4.1E-07</b>	2.2	<b>5.9E-07</b>	3.2	<b>6.3E-07</b>	3.5	<b>4.6E-07</b>	2.6	<b>5.5E-07</b>	3.3
Edad*sexo	<b>0.006</b>	6.5	<b>0.006</b>	6.8	<b>0.003</b>	3.1	<b>0.004</b>	5.0	<b>0.003</b>	3.8
Constante	<b>-3.148</b>	-2.9	<b>-2.854</b>	-32.8	<b>-3.035</b>	-35.5	<b>-2.879</b>	-34.6	<b>-2.865</b>	-35.7
Pseudo R <sup>2</sup>	0.073		0.080		0.072		0.076		0.074	
Región de área común	[.04076071, .85300305]		[.04316973, .80913491]		[.04849855, .84587824]		[.07463188, .86180901]		[.04941797, .86054823]	

Nota: Los coeficientes en negritas son estadísticamente significativos al nivel del 5% o menos.

precisamente el impacto de la informalidad. Para una mejor explicación al respecto, véase Juárez (2006). En cualquier caso, nuestros resultados son robustos a un cambio en la definición de informalidad.

<sup>14</sup> Además de estas especificaciones, se evaluaron otras que no resultaron adecuadas. Por ejemplo, en un caso se utilizó la especificación que utilizan Pratap y Quintin (2005), en donde el modelo probit incluye como variables independientes las siguientes: edad, sexo, una variable indicadora que toma el valor de 1, si otro miembro de la familia es empleado del sector formal y variables dummies del tamaño del establecimiento y la educación. No obstante, esta especificación no cumple con la propiedad de balanceo.

Cuadro 2. *Continuación*

Variable	2000		2001		2002		2003		2004	
	Coef.	t	Coef.	t	Coef.	T	Coef.	t	Coef.	t
Familiar (Formal=1)	<b>0.509</b>	24.6	<b>0.499</b>	25.5	<b>0.454</b>	23.9	<b>0.383</b>	20.5	<b>0.450</b>	21.7
Estado civil (Casado=1)	<b>0.101</b>	11.0	<b>0.044</b>	5.1	<b>0.080</b>	8.8	<b>0.094</b>	9.8	<b>0.077</b>	6.4
Sexo (Hombre=1)	<b>-0.209</b>	-8.8	<b>-0.199</b>	-9.0	<b>-0.135</b>	-5.6	<b>-0.173</b>	-6.7	<b>-0.123</b>	-3.8
Edad	<b>0.102</b>	31.4	<b>0.107</b>	34.5	<b>0.101</b>	30.0	<b>0.107</b>	29.5	<b>0.100</b>	21.9
Edad <sup>2</sup>	<b>-0.001</b>	-34.7	<b>-0.001</b>	-40.5	<b>-0.001</b>	-32.2	<b>-0.001</b>	-31.5	<b>-0.001</b>	-23.8
Escolaridad	<b>0.226</b>	23.2	<b>0.185</b>	19.4	<b>0.234</b>	22.7	<b>0.233</b>	20.8	<b>0.198</b>	14.2
Escolar <sup>2</sup>	<b>-0.006</b>	-19.4	<b>-0.004</b>	-15.2	<b>-0.006</b>	-18.5	<b>-0.005</b>	-15.8	<b>-0.004</b>	-11.1
Escoedad	<b>-0.002</b>	-7.8	<b>-0.001</b>	-4.5	<b>-0.002</b>	-8.5	<b>-0.002</b>	-8.0	<b>-0.001</b>	-3.5
Escoedad <sup>2</sup>	<b>7.1E-07</b>	4.5	<b>2.0E-07</b>	1.3	<b>8.1E-07</b>	5.1	<b>7.0E-07</b>	4.1	<b>2.0E-07</b>	1.0
Edadsexo	<b>0.003</b>	4.9	<b>0.007</b>	10.9	0.001	1.4	<b>0.002</b>	2.6	0.001	1.2
Constante	<b>-3.042</b>	-38.3	<b>-3.136</b>	-40.7	<b>-3.084</b>	-36.6	<b>-3.218</b>	-35.1	<b>-3.154</b>	-27.4
Pseudo R <sup>2</sup>	0.071		0.075		0.068		0.074		0.075	
Región de área común	[.0604379, .85376375]		[.02329611, .83854838]		[.07863285, .8499514]		[.05058889, .85830148]		[.07385831, .84742091]	

Nota: Los coeficientes en negritas son estadísticamente significativos al nivel del 5% o menos.

Cuadro 3. **Estimación Probit (Segunda especificación)**  
**Variable dependiente: FORMAL**

Variable	1995		1996		1997		1998		1999	
	Coef.	t	Coef.	t	Coef.	T	Coef.	t	Coef.	T
Familiar (Formal=1)	<b>0.268</b>	20.2	<b>0.268</b>	21.3	<b>0.264</b>	21.4	<b>0.324</b>	24.9	<b>0.297</b>	22.8
Sexo (Hombre=1)	<b>-0.389</b>	-13.6	<b>-0.361</b>	-13.1	<b>-0.231</b>	-8.6	<b>-0.270</b>	-10.3	<b>-0.218</b>	-8.8
Edad	<b>0.115</b>	30.9	<b>0.115</b>	31.7	<b>0.121</b>	34.2	<b>0.114</b>	33.4	<b>0.108</b>	32.9
Edad <sup>2</sup>	<b>-0.001</b>	-33.4	<b>-0.001</b>	-35.2	<b>-0.001</b>	-37.4	<b>-0.001</b>	-37.5	<b>-0.001</b>	-35.6
Escolaridad	<b>0.229</b>	20.2	<b>0.228</b>	20.6	<b>0.230</b>	21.2	<b>0.216</b>	20.4	<b>0.234</b>	23.0
Escolar <sup>2</sup>	<b>-0.006</b>	-17.8	<b>-0.006</b>	-19.0	<b>-0.006</b>	-19.5	<b>-0.006</b>	-17.5	<b>-0.006</b>	-19.3
Escoedad	<b>-0.002</b>	-6.1	<b>-0.002</b>	-6.2	<b>-0.002</b>	-6.5	<b>-0.002</b>	-5.7	<b>-0.002</b>	-7.2
Escoedad <sup>2</sup>	<b>4.1E-07</b>	2.1	<b>5.9E-07</b>	3.2	<b>6.2E-07</b>	3.4	<b>4.5E-07</b>	2.5	<b>5.5E-07</b>	3.2
Edadsexo	<b>0.007</b>	7.8	<b>0.007</b>	8.5	<b>0.004</b>	4.6	<b>0.005</b>	6.5	<b>0.004</b>	5.2
Constante	<b>-2.945</b>	-33.3	<b>-2.970</b>	-34.3	<b>-3.127</b>	-36.8	<b>-2.971</b>	-35.9	<b>-2.956</b>	-37.1
Pseudo R <sup>2</sup>	0.0721		0.0702		0.0716		0.0751		0.0733	
Región de área común	[.04198033, .84335144]		[.04913495, .82876982]		[.05095912, .83469529]		[.07634869, .85134771]		[.05170092, .85022494]	

Nota: Los coeficientes en negritas son estadísticamente significativos al nivel del 5% o menos.

Cuadro 3. *Continuación*

Variable	2000		2001		2002		2003		2004	
	Coef.	t	Coef.	t	Coef.	T	Coef.	t	Coef.	t
Familiar (Formal=1)	<b>0.509</b>	24.6	<b>0.499</b>	25.6	<b>0.455</b>	24.0	<b>0.383</b>	20.5	<b>0.450</b>	21.7
Sexo (Hombre=1)	<b>-0.219</b>	-9.3	<b>-0.204</b>	-9.2	<b>-0.142</b>	-5.8	<b>-0.181</b>	-7.0	<b>-0.131</b>	-4.1
Edad	<b>0.110</b>	34.6	<b>0.110</b>	36.4	<b>0.107</b>	32.5	<b>0.114</b>	32.0	<b>0.106</b>	23.5
Edad <sup>2</sup>	<b>-0.001</b>	-39.0	<b>-0.001</b>	-43.3	<b>-0.001</b>	-35.5	<b>-0.001</b>	-35.0	<b>-0.001</b>	-26.1
Escolaridad	<b>0.224</b>	23.0	<b>0.184</b>	19.3	<b>0.233</b>	22.6	<b>0.232</b>	20.8	<b>0.197</b>	14.2
Escolar <sup>2</sup>	<b>-0.006</b>	-19.5	<b>-0.004</b>	-15.3	<b>-0.006</b>	-18.7	<b>-0.005</b>	-16.0	<b>-0.004</b>	-11.2
Escoedad	<b>-0.002</b>	-7.5	<b>-0.001</b>	-4.4	<b>-0.002</b>	-8.4	<b>-0.002</b>	-7.8	<b>-0.001</b>	-3.4
Escoedad <sup>2</sup>	<b>7.0E-07</b>	4.4	2.0E-07	1.3	<b>8.1E-07</b>	5.1	<b>7.1E-07</b>	4.2	2.0E-07	1.0
Edadsexo	<b>0.004</b>	6.4	<b>0.007</b>	11.6	<b>0.002</b>	2.5	<b>0.003</b>	3.8	<b>0.002</b>	2.1
Constante	<b>-3.130</b>	-39.6	<b>-3.172</b>	-41.4	<b>-3.155</b>	-37.6	<b>-3.301</b>	-36.2	<b>-3.217</b>	-28.1
Pseudo R <sup>2</sup>	0.0703		0.0746		0.0670		0.0732		0.0751	
Región de área común	[.06357105, .84306864]		[.02312913, .83713415]		[.08032886, .84166428]		[.04996137, .8481712]		[.07449041, .83911691]	

Nota: Los coeficientes en negritas son estadísticamente significativos al nivel del 5% o menos.

Los resultados obtenidos con el modelo *Probit* son, en su gran mayoría, significativos y con los signos esperados. Por ejemplo, los resultados muestran que la probabilidad de participar en el sector formal es más alta para personas del sexo femenino, de mayor edad, casadas y con un mayor nivel de escolaridad. Algunas de estas variables ya se habían mencionado anteriormente como factores que tienden a caracterizar a los trabajadores del sector formal (véase cuadro 1).

Un resultado un tanto sorpresivo es el efecto positivo que resulta de tener un familiar que trabaja en el sector formal, sobre la probabilidad de trabajar en este mismo sector. Este resultado es bastante robusto a lo largo del tiempo y su signo positivo es un tanto inesperado, considerando que algunos trabajos anteriores sugerían la existencia de un comportamiento estratégico entre los individuos de una familia, el cual los llevaría a buscar una cierta diversificación en sus actividades laborales. Este es el caso, por ejemplo, planteado en el modelo de Galiani y Weinschelbaum (2008), en donde la participación en el mercado formal por parte de trabajadores semi-calificados disminuye cuando ya existe un miembro del hogar que labora en el sector formal. En última instancia, este resultado puede verse como una evidencia más en contra de los modelos que suponen la existencia de integración y libre movilidad entre los mercados laborales.<sup>15</sup>

<sup>15</sup> En cualquier caso, este resultado no es crucial para los resultados de este trabajo. El anexo 2 muestra los resultados del Efecto de Tratamiento Promedio que se hubieran obtenido si optamos por excluir la variable *familiar* de la estimación del *propensity score*. Como allí se menciona, los efectos de tratamiento encontrados son prácticamente los mismos. Este resultado quizá pueda interpretarse como una confirmación de los resultados de Zhao (2008), quien enfatiza la robustez de los métodos que utilizan el *propensity score*, siempre y cuando se cumplan los supuestos relevantes.

### ***Efectos de Tratamiento Promedio (Brechas Salariales)***

Con los resultados de las estimaciones de los cuadros 2 y 3 es posible estimar la propensión a participar en el sector formal de cada uno de los individuos ( $p_i$ ) y el efecto de tratamiento promedio definido en la ecuación (3.5). Este efecto se puede calcular a través de los cuatro métodos, que están basados en el *propensity score matching*, que se menciona y describen con mayor detalle, en el anexo 1.

Antes de realizar este ejercicio, cabe hacer algunas precisiones. Primero, debe notarse que la variable de interés es el salario por hora trabajada. Por ello, para estimar esta variable se multiplicó el salario mensual después de impuestos que se reporta en la ENEU, por el número de meses (12) y se dividió por el número en el año (52), para obtener una estimación del salario semanal en el año de referencia.

Posteriormente, este valor se dividió entre el número de horas que el individuo reportó como trabajadas en una semana, para obtener una estimación del salario por hora trabajada durante el periodo. Segundo, al haber definido el efecto de tratamiento como la brecha del logaritmo de los salarios por hora en ambos sectores, el efecto de tratamiento en realidad se refiere a una proporción, la cual puede interpretarse directamente como el premio salarial del sector formal, en relación con el sector informal.

En el cuadro 4, se muestra la estimación por 4 diferentes métodos, del premio salarial promedio en el sector formal, para cada uno de los años de la muestra.<sup>16</sup> Como ya se mencionó, la variable dependiente es la diferencia en el logaritmo del salario por hora de cada uno de los sectores. En las estimaciones se utilizó la opción de área común (*common support*), lo cual implica que sólo se utilizan aquellas observaciones cuya propensión a participar en el sector formal cae en la región en donde existen suficientes datos de ambos sectores para hacer las comparaciones apropiadas. Los cuadros 4a y 4b muestran los resultados obtenidos con la primera y la segunda especificación, respectivamente. Con estos dos grupos de resultados, se realizan las gráficas 7a y 7b.<sup>17</sup>

El cuadro 4 y la gráfica 7 muestran que, independientemente de la especificación utilizada para la estimación del *propensity score* y del método de estimación del efecto de tratamiento, el efecto estimado es positivo y muy significativo, lo cual sugiere la existencia de un premio salarial al trabajo

---

<sup>16</sup> En la estimación que utiliza el *Kernel*, se utilizó una distribución normal y un ancho de banda igual a 0.06. En la estimación de *Radius Matching*, ésta se hizo con un radio de 0.1.

<sup>17</sup> Las estimaciones fueron realizadas con el programa Stata, y se utilizó el módulo desarrollado por Becker e Ichino (2002).

## 24 Ensayos

formal. Es decir, que individuos que trabajan en el sector formal obtienen, en promedio, un salario por hora más elevado que el que reciben individuos con características similares pero que laboran en el sector informal.

Este resultado contradice las implicaciones de la hipótesis de los mercados laborales integrados, y favorece la hipótesis de segmentación en el mercado laboral en México.

**Cuadro 4<sup>a</sup>. Average treatment effect  
(Primera especificación)**

Año	Nearest Neighbor		Radius		Kernel		Stratification	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t
1995	0.118	16.26	0.170	26.29	0.150	24.11	0.117	17.90
1996	0.142	20.27	0.205	32.96	0.182	37.33	0.148	23.62
1997	0.178	25.98	0.238	38.65	0.221	236.43	0.179	29.76
1998	0.176	26.64	0.247	41.94	0.223	35.62	0.178	30.40
1999	0.175	29.10	0.247	46.10	0.223	60.96	0.180	33.05
2000	0.145	25.40	0.218	41.59	0.198	51.27	0.155	30.10
2001	0.134	24.28	0.204	40.66	0.180	84.32	0.136	27.12
2002	0.142	26.58	0.205	42.17	0.185	26.81	0.144	29.66
2003	0.137	24.04	0.211	41.73	0.180	63.59	0.141	27.34
2004	0.115	15.80	0.189	29.95	0.156	23.29	0.119	18.49

Nota: La variable dependiente es el logaritmo del salario por hora. Los coeficientes en negritas son estadísticamente significativos al nivel del 5% o menos.

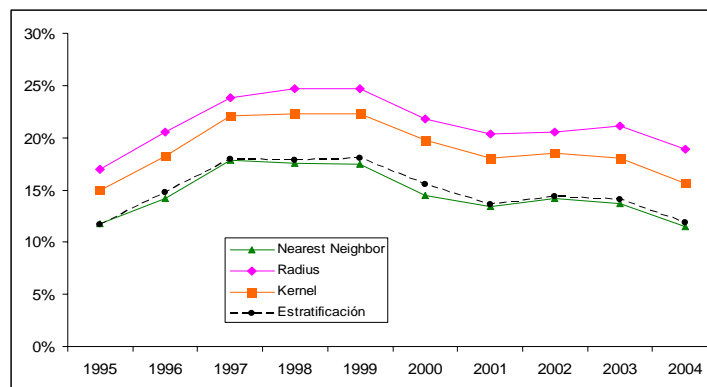
**Cuadro 4b. Average treatment effect  
(Segunda especificación)**

Año	Nearest Neighbor		Radius		Kernel		Stratification	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t
1995	0.119	16.73	0.171	26.45	0.153	38.90	0.124	18.82
1996	0.142	20.88	0.207	33.16	0.187	30.96	0.153	24.21
1997	0.18	27.02	0.239	38.84	0.225	40.53	0.184	30.45
1998	0.179	27.58	0.248	42.19	0.227	58.28	0.184	31.45
1999	0.177	30.15	0.249	46.38	0.227	37.92	0.183	33.62
2000	0.148	26.42	0.220	41.96	0.203	245.93	0.159	30.69
2001	0.136	25.11	0.207	41.19	0.184	53.34	0.137	27.28
2002	0.143	27.31	0.207	42.56	0.188	39.50	0.147	30.05
2003	0.137	24.70	0.211	41.73	0.183	49.72	0.143	27.66
2004	0.117	16.70	0.191	30.16	0.159	547.52	0.124	19.27

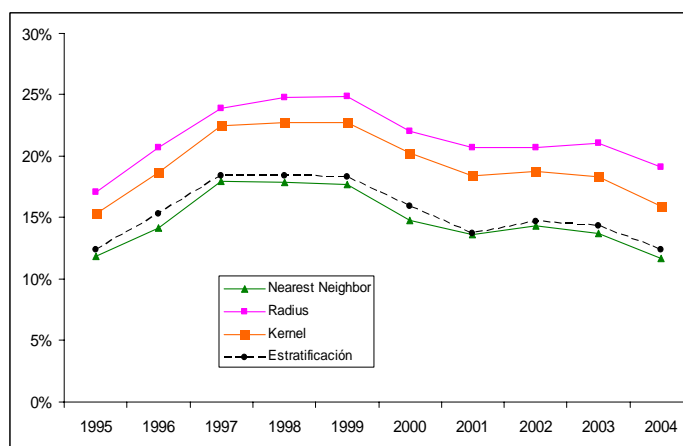
Nota: La variable dependiente es el logaritmo del salario por hora. Los coeficientes en negritas son estadísticamente significativos al nivel del 5% o menos.



**Gráfica 7. Premio salarial promedio en el sector formal  
(primera especificación)**



**Gráfica 7b. Premio salarial promedio en el sector formal  
(segunda especificación)**



En términos de la magnitud del efecto de tratamiento o de la brecha salarial entre sectores, los resultados obtenidos dan muestra de ser muy consistentes respecto de los métodos del “vecino más cercano” (*nearest neighbor*) y el de estratificación; mientras que los métodos que utilizan un radio y un kernel como método de comparación, tienden a estimar efectos de tratamiento (o brechas salariales) ligeramente más elevados. En todos los casos, también el efecto de tratamiento estimado en 2004 es prácticamente el mismo, al de inicio del periodo (1995); por lo que no puede concluirse que el

premio salarial al sector formal haya mostrado una tendencia única en años recientes, que pudiera estar asociada al comportamiento de los recursos destinados a la política social.

En todos los casos, sin embargo, los resultados muestran un comportamiento muy similar entre sí, con un ligero aumento entre 1995 y 1997, una estabilización entre 1997 y 1999, y una reducción paulatina a partir de entonces y hasta 2004. Así, el tamaño estimado de la prima salarial al sector formal fluctúa entre 12% y 17% en 1995, alcanza un máximo que fluctúa entre 18% y 25% en 1998-99, y termina en un rango de entre 12% y 17% en 2004.<sup>18</sup> Estos resultados implican que el mercado de trabajo de la economía mexicana, a diferencia de lo que sugieren o suponen algunos trabajos previos, no es competitivo y que el mercado laboral mexicano está segmentado.<sup>19</sup>

Por otro lado, una comparación simple entre la brecha salarial después de haber controlado por diferencias observables entre los individuos que laboran en uno u otro sector y la brecha salarial no condicional (gráfica 6), muestra la existencia de un patrón cíclico notablemente similar, y sugiere que entre 35% y 45% de la brecha salarial observada en los ingresos de los trabajadores formales e informales, se explica por la segmentación del mercado laboral, mientras que el resto es explicable por las diferencias observables entre los individuos de uno y otro sector en variables asociadas a la educación, el género, la edad, la experiencia, etc.

---

<sup>18</sup> Con el objetivo de verificar que los resultados obtenidos fueran robustos a cambios en la muestra y que no estuvieran influidos por trabajadores de tiempo parcial, también se realizaron estimaciones eliminando de la muestra a quienes declararon trabajar menos de 10 horas a la semana o menos de 20 horas. En ambos casos, se obtuvieron resultados muy similares a los que se reportan en la gráfica 7.

<sup>19</sup> Cabe señalar que los resultados de este ejercicio son muy diferentes a los obtenidos por Levy (2008), en un ejercicio aparentemente similar y que también utiliza la técnica de Efectos de Tratamiento y de *matching* (véase el cap. 5, cuadros 5.12 y 5.13). Según estos resultados, los trabajadores de bajos salarios no obtienen ingresos estadísticamente diferentes al cambiar de sector, mientras que los trabajadores de altos ingresos que pasan del sector formal al informal reciben un ingreso más elevado al cambiar de sector, pero no ocurre así, si lo hacen en la dirección contraria (excepto en el caso de las mujeres). Dicho ejercicio es muy diferente al nuestro, no sólo en términos de su muestra (la cual sólo utiliza información para el periodo 2005-2006) sino, sobre todo, en términos conceptuales, ya que el ejercicio de Levy (2008) sólo analiza el efecto en los salarios de aquellos trabajadores que cambiaron de un sector al otro en el periodo analizado. Es claro que este criterio para seleccionar la muestra, por sí mismo, puede indicar un sesgo importante en términos de características no observables, además de que elimina de la comparación a la gran mayoría de los trabajadores de ambos sectores que no cambiaron de empleo, ni de sector.

Finalmente, debe señalarse que la evolución temporal del premio salarial al sector formal quizá pueda explicarse como resultado del comportamiento cíclico de la economía. Debe recordarse que los años en los que el premio salarial al sector formal aumentó -de 1995 a 1997 y de 2001 a 2002-, coinciden con periodos de fuertes presiones en el mercado laboral, asociadas a la crisis macroeconómica y financiera de 1994/95 en México y, a la recesión en México y en Estados Unidos de 2001.

En suma, podemos concluir que no existe evidencia empírica a favor de la hipótesis de que la política social esté fomentando la informalidad en México. Esto último puede explicarse como el resultado natural de que los mercados laborales en México estén segmentados y no integrados, como lo presupone dicha hipótesis. Es por ello que las razones de la relativamente alta informalidad en México quizá deban buscarse en otra parte y no necesariamente en el diseño de la política social.

### **Resumen y Conclusiones**

En un estudio reciente (Levy, 2008), se plantea la hipótesis de que el aumento en los recursos destinados a ciertos programas sociales tiende a fomentar la informalidad. Este argumento, sin embargo, parte de suponer que el mercado de trabajo es competitivo y que los trabajadores se pueden mover libremente entre el sector formal e informal de la economía. En el presente trabajo, se analiza empíricamente la validez de ese argumento para el caso de México, durante el periodo 1995-2004.

Entre 1998 y 2007, los subsidios destinados para la protección social crecieron en 110%, mientras que los recursos destinados a la seguridad social lo hicieron en alrededor de 40%. Si la hipótesis de que los programas sociales tienden a fomentar la informalidad en México fuera cierta, se habría esperado observar, al menos durante una parte del periodo analizado, un incremento en la tasa de informalidad de la economía. Por otro lado, si la hipótesis de que los mercados laborales están integrados, fuera correcta, se esperaría observar la existencia de un premio salarial a favor de los trabajadores informales; es decir, que individuos que laboran en este sector ganarían más que individuos con características similares, pero que trabajan en el sector formal de la economía.

Los resultados obtenidos en este trabajo demuestran que ninguna de estas dos situaciones se ha cumplido en el caso de México. Por una parte, la tasa de informalidad en el país, entre los trabajadores asalariados, se mantuvo bastante estable a lo largo de todo el periodo analizado e independientemente de la definición de informalidad adoptada. Por otro lado, los resultados que se basaron en la técnica del *propensity score*

*matching*, muestran que existe evidencia muy consistente en el sentido de que existe un premio salarial al trabajo en el sector formal de la economía, por lo que los mercados de trabajo en México no parece estén integrados, como lo supone la hipótesis en cuestión. Por lo tanto, la evidencia que se ha obtenido en este trabajo, permite afirmar que: los aumentos recientes en los programas sociales en México no han sido una causa de la informalidad en el país.

## Referencias

- Amaral, P y E. Quintin (2006); "A Competitive Model for the Informal Sector", *Journal of Monetary Economics*, 53, 1541-53.
- Becker, S. y A. Ichino (2002). "Estimation of Average Treatment Effects Based on Propensity Scores" *The Stata Journal*, Vol. 2(4), 358-377
- De Soto, Hernando (1987), *El Otro Sendero: la Revolución Informal*, Editorial Diana, 317 p.
- Dehejia, R. y S. Wahba (1999). "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs", *Journal of the American Statistical Association*, Vol 94, No. 488, 1053-1062.
- Dehejia, R. y S. Wahba (2002). "Propensity Score Matching Methods for Nonexperimental Causal Studies", *Review of Economics and Statistics*, Vo. 84, 151-161.
- Galiani, S. y F. Weinschelbaum (2008), "Modeling Informality Formally: Households and Firms", mimeo, Julio.
- Gasparini, L., y L. Tornarolli (2007), "Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata.", Documento de Trabajo, no. 46, CEDLAS, Febrero.
- Gong, X., y A. van Soest (2002), "Wage Differentials and Mobility in the Urban Labour Market: A Panel Data Analysis for Mexico", *Labour Economics*, 9, 513-29.
- Harris, J. y M. Todaro (1970), "Migration, Unemployment and Development: A Two-Sector Analysis," *American Economic Review* 60, 126-142.
- Juárez, L. (2006), "Wage differentials between formal and informal salaried workers in Mexico", mimeo, ITAM.
- Levy, S. (2007), "¿Pueden los programas sociales disminuir la productividad y el crecimiento económico? Una hipótesis para México", *El Trimestre Económico*, Vol. LXXIV, No. 3, 491-540.
- (2008), *Good Intentions, Bad Outcomes: Social Policy, Informality and Economic Growth in Mexico*, Brookings Institution Press, Washington, D.C.
- Maloney, W. (1999), "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from sectoral Transitions in Mexico," *The World Bank Economic Review* Vol. 13, No. 2, 275-302.

- (2004), "Informality Revisited," *World Bank Policy Research Working Paper*, No. 2965.
- Marcouiller, D., V. Ruiz de Castilla, V. y C. Woodruff (1997), "Formal Measures of the Informal Sector Wage Gap in Mexico, El Salvador, and Peru," *Economic Development and Cultural Change* Vol. 45, No. 2, 367-92.
- Navarro, S. y P. Schrimpf (2004), "The importance of being formal: testing for segmentation in the Mexican labor market," mimeo.
- Perry, G., W. F. Maloney, O. S. Arias, P. Fajnzylber, A. Mason y J. Saavedra (2007), *Informalidad: Escape y Exclusión*, Banco Mundial, Washington, D.C.
- Pradhan, M. y A. van Soest (1995), "Formal and informal sector employment in urban areas of Bolivia", *Labour Economics*, 2, 275-297.
- Pratap, S. y E. Quintin (2006), "Are Labor Market Segmented in Developing Countries? A Semiparametric Approach," *European Economic Review*, 50-7, October, 1817-41.
- Rauch, J. (1991); "Modeling the Informal Sector Formally", *Journal of Development Economics*, 35, pp. 33-47.
- Rosenbaum, P. y D. Rubin (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika* 70(1), 41-55.
- Rosenbaum, P. y D. Rubin (1984), "Reducing Bias in Observational Studies using Sub Classification on the Propensity Score," *Journal of the American Statistical Association* 79, 516-524.
- Rubin, D. (1974), "Estimating Causal effects of Treatment in Randomized and Nonrandomized Studies," *Journal of Education Psychology* Vol. 66, 688-701.
- Zhao, Zhong (2008), "Sensitivity of Propensity Score Methods to the Specifications", *Economics Letters*, 98, 309-19. Anexo 1: Métodos de Estimación del Efecto de Tratamiento con el *Propensity Score*.

## **Anexo 1 Métodos de apareamiento**

### ***Nearest Neighbor y Radius Matching***

El método de *Nearest Neighbor* consiste en tomar a cada individuo del sector formal y comparar el logaritmo de su salario con el del individuo del sector informal cuyo *propensity score* sea el más cercano. El ATE se obtiene mediante el promedio de las brechas salariales entre todos los pares de individuos que se puedan construir. Por otro lado, el método de *Radius Matching* en lugar de utilizar al *propensity score* del individuo más cercano, toma a todos aquellos individuos del grupo de control cuyo *propensity score* cae dentro de una vecindad previamente especificada.

La fórmula de ambos tipos de estimadores se puede escribir de la siguiente forma:

$$\begin{aligned}\alpha^{NN,R} &= \frac{1}{N^F} \sum_{k \in F} (w_{f_k} - \sum_{j \in I(k)} \theta_{kj} w_{i_j}) \\ &= \frac{1}{N^F} \left[ \sum_{k \in F} w_{f_k} - \sum_{k \in F} \sum_{j \in I(k)} \theta_{kj} w_{i_j} \right]\end{aligned}\quad (\text{A.1})$$

donde  $N^F$  es el número de trabajadores formales en la muestra,  $I(k)$  representa el conjunto de unidades del sector informal emparejadas a la unidad  $k$ -ésima del sector formal, con un valor estimado del propensity score  $p_k$ , y las ponderaciones  $\theta_{kj} = \frac{1}{N_k^I}$  si  $j \in I(k)$  y  $\theta_{ij} = 0$  de otra forma.

Donde  $I(k) = \min_j \|p_k - p_j\|$  en el método *Nearest Matching* e

$I(k) = \{j \mid \|p_k - p_j\| < r\}$  en el método *Radius Matching*, con un radio igual a  $r$ .

### Kernel Matching

Con el método de *Kernel*, las observaciones de todos los individuos del sector formal son “emparejadas” con un promedio ponderado de las de los trabajadores informales, con ponderaciones que son inversamente proporcionales a la distancia entre los *propensity scores* de formales e informales. El estimador está dado por:

$$\alpha^K = \frac{1}{N^F} \sum_{k \in F} \left\{ w_{f_k} - \frac{\sum_{j \in I} w_{i_j} K\left(\frac{p_j - p_k}{h_n}\right)}{\sum_{q \in I} K\left(\frac{p_q - p_k}{h_n}\right)} \right\} \quad (\text{A.2})$$

donde  $K(\cdot)$  es una función kernel y  $h_n$  es el ancho de banda.

### **Stratification Matching**

El método *Stratification matching* consiste en dividir el rango de variación de los *propensity scores* en intervalos, de forma tal que dentro de cada uno, las observaciones de los trabajadores formales e informales tengan -en promedio- el mismo *propensity score*. El efecto del tratamiento en cada intervalo se puede definir por medio de la siguiente expresión:

$$\alpha_q^S = \frac{\sum_{k \in S(q)} w_{f_k}}{N_q^F} - \frac{\sum_{j \in S(q)} w_{i_j}}{N_q^I} \quad (\text{A.3})$$

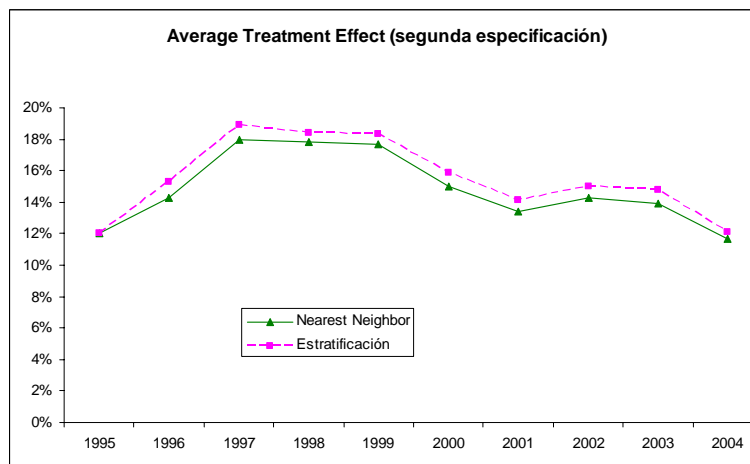
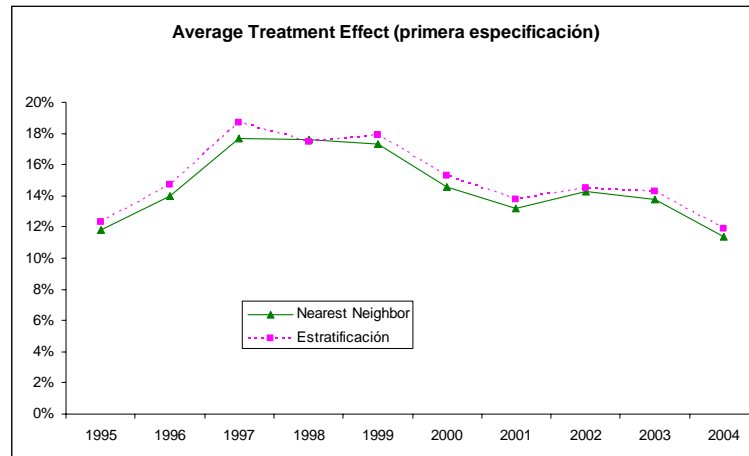
donde  $S(q)$  es el grupo de unidades en el estrato  $q$ .  $N_q^F$  y  $N_q^I$  son las cantidades de unidades de formales e informales en el estrato  $q$ .

El estimador del ATE se obtiene de la forma siguiente:

$$\alpha^S = \sum_{q=1}^Q \alpha_q^S \frac{\sum_{k \in S(q)} Formal_k}{\sum_{\forall k} Formal_k} \quad (\text{A.4})$$

### **Anexo 2. Estimaciones sin incluir la variable *Familiar***

Si no se incluyera la variable familiar, las estimaciones de los efectos de tratamiento utilizando las metodologías de *Nearest Neighbor* y *Stratification* para las dos estimaciones mencionadas en el texto, serían las siguientes:



Como se puede ver, no hay diferencias significativas con respecto a los resultados reportados en el texto.